Retroposed Elements and Their Flanking Regions Resolve the Evolutionary History of Xenarthran Mammals (Armadillos, Anteaters, and Sloths)

Maren Möller-Krull,*¹ Frédéric Delsuc,†¹ Gennady Churakov,* Claudia Marker,* Mariella Superina,‡ Jürgen Brosius,* Emmanuel J. P. Douzery,†² and Jürgen Schmitz*²

*Institute of Experimental Pathology (ZMBE), University of Münster, Münster, Germany; †Institut des Sciences de l'Evolution de Montpellier (ISE-M), UMR5554-CNRS, Université Montpellier II, Montpellier, France; and ‡Department of Biological Sciences, University of New Orleans

Armadillos, anteaters, and sloths (Order Xenarthra) comprise 1 of the 4 major clades of placental mammals. Isolated in South America from the other continental landmasses, xenarthrans diverged over a period of about 65 Myr, leaving more than 200 extinct genera and only 31 living species. The presence of both ancestral and highly derived anatomical features has made morphoanatomical analyses of the xenarthran evolutionary history difficult, and previous molecular analyses failed to resolve the relationships within armadillo subfamilies. We investigated the presence/absence patterns of retroposons from \sim 7,400 genomic loci, identifying 35 phylogenetically informative elements and an additional 39 informative rare genomic changes (RGCs). DAS-short interspersed elements (SINEs), previously described only in the Dasypus novemcinctus genome, were found in all living armadillo genera, including the previously unsampled Chlamyphorus, but were noticeably absent in sloths. The presence/absence patterns of the phylogenetically informative retroposed elements and other RGCs were then compared with data from the DNA sequences of the more than 12-kb flanking regions of these retroposons. Together, these data provide the first fully resolved genus tree of xenarthrans. Interestingly, multiple evidence supports the grouping of *Chaetophractus* and *Zaedyus* as a sister group to *Euphractus* within Euphractinae, an association that was not previously demonstrated. Also, flanking sequence analyses favor a close phylogenetic relationship between Cabassous and Tolypeutes within Tolypeutinae. Finally, the phylogenetic position of the subfamily Chlamyphorinae is resolved by the noncoding sequence data set as the sister group of Tolypeutinae. The data provide a stable phylogenetic framework for further evolutionary investigations of xenarthrans and important information for defining conservation priorities to save the diversity of one of the most curious groups of mammals.

Introduction

Sloths, anteaters, and armadillos are among the most morphologically specialized mammals. These 3 lineages belong to Xenarthra, a mammalian order whose monophyly is undisputed (Engelmann 1985; Van Dijk et al. 1999). On molecular grounds, xenarthrans constitute 1 of the 4 major placental clades (Madsen et al. 2001; Murphy, Eizirik, Johnson, et al. 2001; Delsuc et al. 2002). Their evolutionary radiation began about 65 MYA while South America was becoming isolated from other continental land masses (Patterson and Pascual 1972). A total of 31 living species of xenarthrans (21 armadillos, 4 anteaters, and 6 sloths) have been described (Wetzel 1985; Vizcaíno 1995; Anderson and Handley 2001). Today's limited xenarthran species number is in sharp contrast with their past diversity as more than 200 fossil genera are recognized (McKenna and Bell 1997). The most impressive xenarthrans, such as giant ground sloths and glyptodonts, disappeared at the end of the last Pleistocene glaciations (about 10,000 years ago; Patterson and Pascual 1972). Remnants of this ancient evolutionary radiation, most extant xenarthran species, are, nevertheless, poorly studied and uncertainties remain concerning their conservation status (da Fonseca and Aguiar 2004).

As curious as xenarthrans are, so disputed is their place in the placental evolutionary tree. Whereas nuclear gene concatenations have been frustratingly inconclusive

¹ M.M.-K. and F.D. equally contributed to this study.

² E.J.P.D. and J.S. share senior authorship.

Key words: DAS-SINE, retroposon, phylogeny, rare genomic changes, maximum likelihood.

E-mail: jueschm@uni-muenster.de.

Mol. Biol. Evol. 24(11):2573–2582. 2007 doi:10.1093/molbev/msm201 Advance Access publication September 19, 2007

© The Author 2007. Published by Oxford University Press on behalf of the Society for Molecular Biology and Evolution. All rights reserved. For permissions, please e-mail: journals.permissions@oxfordjournals.org

(Murphy, Eizirik, O'Brien, et al. 2001; Delsuc et al. 2002; Amrine-Madsen et al. 2003; Nikolaev et al. 2007), the classical morphology-based Epitheria hypothesis, placing Xenarthra as the sister group of all placental orders (McKenna 1975), has found some support from retroposed elements (Kriegs et al. 2006). However, Murphy et al. (2007) have recently found some rare insertion/deletion events that represent contradictory evidence in support of a sister group relationship between Xenarthra and Afrotheria instead.

Living xenarthrans are divided into 3 phylogenetically distinct groups, the arboreal sloths (Folivora: Megalonychidae and Bradypodidae), the toothless anteaters (Vermilingua: Myrmecophagidae), and the shelled armadillos (Cingulata: Dasypodidae). Folivora and Vermilingua form the suborder Pilosa for which there is unequivocal support from both morphological (Engelmann 1985; Patterson et al. 1992) and molecular (Delsuc et al. 2001, 2002) data.

The phylogenetic relationships and taxonomy of the living pilosans are well resolved (Gaudin and Branham 1998; Delsuc et al. 2001, 2002). Folivora contains the 2 morphologically distinct tree sloth families Megalonychidae (*Choloepus*, 2-toed sloths) and Bradypodidae (*Bradypus*, 3-toed sloths) (Webb 1985; Höss et al. 1996; Greenwood et al. 2001). Among Vermilingua, myological (Reiss 1997) and morphological (Gaudin and Branham 1998) studies supporting an early emergence of the pygmy anteater (*Cyclopes didactylus*) have been confirmed by molecular data (Delsuc et al. 2001, 2002). These studies therefore group together the giant anteater (*Myrmecophaga*) and tamanduas (*Tamandua*).

The armadillos are the oldest and most speciose lineage of xenarthrans. Their phylogeny has been difficult to reconstruct from morphoanatomical data with incongruent results among the rare studies (Engelmann 1985; Patterson et al. 1989; Abrantes and Bergqvist 2006; Gaudin and Wible 2006). Molecular evidence (Delsuc et al. 2002, 2003) has, nevertheless, clearly defined 3 major lineages compatible with the subfamilies of the classification of McKenna and Bell (1997): Dasypodinae (the single genus *Dasypus*), Tolypeutinae (genera *Tolypeutes, Cabassous,* and *Priodontes*), and Euphractinae (genera *Euphractus, Chaetophractus,* and *Zaedyus*). A close relationship between Tolypeutinae and Euphractinae was also strongly supported, with Dasypodinae emerging first within Cingulata (Delsuc et al. 2001, 2002, 2003). These relationships are congruent with the study of spermatozoa (Cetica et al. 1998) but contradict morphological studies (Engelmann 1985; Gaudin and Wible 2006).

However, even molecular data have failed to resolve relationships within the subfamilies Tolypeutinae and Euphractinae, with contradictory results obtained from mitochondrial versus nuclear genes suggesting that speciation events occurred very rapidly within each of these 2 armadillo lineages (Delsuc et al. 2003). Within tolypeutines, a close relationship between Cabassous and Priodontes would be consistent with their classification in the tribe Priodontini based on their similar morphologies (McKenna and Bell 1997; Abrantes and Bergqvist 2006; Gaudin and Wible 2006) and a study of their spermatozoa (Cetica et al. 1998). Within euphractines, the grouping of *Euphractus* and Chaetophractus to the exclusion of Zaedyus would be congruent with a recent cladistic study of craniodental characters (Gaudin and Wible 2006), whereas another cladistic study including also postcranial characters rather supports a close relationship between Chaetophractus and Zaedyus (Abrantes and Bergqvist 2006). Uncertainty remains in morphological evidence because of the highly derived morphology of *Tolypeutes* within Tolypeutinae and the very similar morphologies exhibited by extant members of Euphractinae.

At present, the only armadillo subfamily for which no molecular data is available is Chlamyphorinae (fairy armadillos). It includes only 2 living species classified in the genus Chlamyphorus: Chlamyphorus truncatus (pink fairy armadillo) and Chlamyphorus retusus (chacoan or greater fairy armadillo). These 2 species are morphologically specialized to a subterranean lifestyle with notably reduced eyes and reinforced forearms supporting enlarged digging claws. Some cladistic analyses of morphological characters including fossils have supported a close relationship of fairy armadillos with Zaedvus (Engelmann 1985; Gaudin and Wible 2006). However, when only extant taxa were considered in an analysis of craniodental characters, Chlamyphorus grouped with Euphractinae (Gaudin and Wible 2006), a result also obtained by Abrantes and Bergqvist (2006) in their morphological study. Yet, the phylogenetic affinities of fairy armadillos have never been tested with molecular data as both species are fairly elusive and rare in the field (Cuéllar 2001; Superina 2006).

The genomic insertion patterns of various retroposed mobile elements provide an extraordinary archive from which to elucidate the evolutionary history of organisms (Shedlock and Okada 2000). Short interspersed elements (SINEs) and other retroposed elements, shared by some taxa but missing from the genomes of others, are highly valuable, essentially homoplasy-free indicators of relatedness (Okada 1991; Nikaido et al. 1999; Salem et al. 2003; Roos et al. 2004; Kriegs et al. 2006; Nishihara et al. 2006), with an extremely large possible number of basically unique character states (Steel and Penny 2005). A specific tRNA-derived retroposed SINE family, called DAS-SINEs, has been discovered in the Dasypus novemcinctus genome (Borodulina and Kramerov 2005; Churakov et al. 2005), with average genomic copy numbers of about 467,000 calculated from the \sim 71% of available genomic information (Ensembl GeneBase version 45.1c). Using a specific strategy to compare nested DAS-SINEs, distinct historical waves of activity of 6 DAS-SINE subfamilies were defined (Churakov et al. 2005). This information was used for the first time in the present study to screen for a maximum variety of young-and respectively old-phylogenetically informative elements to provide the first retroposon presence/ absence support for terminal-and respectively deepsplits in xenarthrans.

Furthermore, low complexity, rare genomic changes (RGCs), such as random insertions or deletions (indels) provide additional capability in molecular systematics (Rokas and Holland 2000). Although simple random indels are of lower reliability than complex retroposon insertions, they have often been shown to be evolutionarily conserved and phylogenetically informative (Murphy et al. 2007). Retroposed elements with their characteristic flanking target site duplications (direct repeats [DRs]), diagnostic mutations, and often specific truncations are exceptionally informative RGCs (Ray et al. 2006) and, compared with simple random indels, are virtually ambiguity free. In the present study, we combined different strategies to find phylogenetically informative retroposed elements and indels to serve as cladistic molecular markers for reconstructing the phylogeny of all xenarthran genera including the previously unsampled genus Chlamyphorus. Moreover, phylogenetic analyses of the flanking regions of these loci enabled us to reconstruct the evolutionary history of xenarthrans using a large data set of non-protein-coding DNA. The resulting phylogenetic tree was used as a framework for mapping the presence/absence of retroposed elements and indels providing independent evidence for further resolving the phylogeny of this major placental clade.

Materials and Methods

Computational Strategies

To identify phylogenetically informative presence/ absence patterns of retroposed elements, we applied 3 different in silico search strategies.

Strategy 1

We downloaded all available *D. novemcinctus* sequences from the National Institutes of Health Intramural Sequencing Center at http://www.ncbi.nlm.nih.gov/entrez/ query.fcgi?db=nucleotide (20.1 million bases) and used the local version of RepeatMasker (Smit and Green at http://www.repeatmasker.org) containing a specific library comprising all DAS-SINE subfamily consensus sequences to scan for DAS-SINEs. All detected DAS-SINEs were extracted along with 250 nt of their flanking sequences (2,023 loci). For outgroup comparison and to construct conserved polymerase chain reaction (PCR) primers, the selected loci were aligned against human genomic sequences using the UCSC Server (http://genome.ucsc.edu/cgibin/hgBlat). Only sequence alignments with more than 80% identity were used to generate conserved PCR primers (80 loci).

For estimating the genomic number of DAS-SINEs in *D. novemcinctus*, we downloaded available sequence information from http://www.ensembl.org/info/data/download. html (Dasypus_novemcinctus.0.ARMA.45.dna.nonchromosomal.fa.gz) and used our specific DAS-SINE library for a local RepeatMasker screen. We extrapolated the number of all detected elements to the expected full size of the genome.

Strategy 2

For finding markers for other xenarthran branches in which DAS-SINEs were possibly not active (deep xenarthran splits), we also searched for other retroposed elements (i.e., long interspersed elements [LINEs] or retroposed small nuclear RNAs [snRNAs] such as spliceosomal U snRNAs) in conserved regions. We downloaded all available trace sequences for D. novemcinctus and Choloepus hoffmanni from ftp://ftp.ncbi.nih.gov/pub/TraceDB/ (4.6 Gb for D. novemcinctus and 4.2 Gb for C. hoffmanni) and detected 5,206 potentially phylogenetic informative loci with embedded L1MB4 (2,609), L1MB5 (625), and L1MB7 (1,891) elements that were potentially active in xenarthrans (Kriegs et al. 2006), as well as 81 snRNA retrogenes. These loci were screened using trace Blast (http:// www.ncbi.nlm.nih.gov/BLAST/mmtrace) in 3 different ways: 1) for the presence of retroposed elements in C. hoffmanni that were absent in D. novemcinctus in order to find potential integrations on the branch leading to sloths, 2) for the presence of elements in D. novemcinctus that were absent in C. hoffmanni in order to find potential integrations specific to Cingulata, and 3) for the presence of elements in both species in order to find potential markers indicating xenarthran monophyly. From the 5,206 candidates, only 32 contained insertions meeting at least one of the above criteria. They were used to generate conserved PCR primers (12 L1MB4, 4 L1MB5, 4 L1MB7, and 12 spliceosomal RNAs).

Strategy 3

To identify randomly inserted xenarthran-specific retroposed elements, we PCR amplified 171 highly conserved mammalian loci that were previously analyzed by Kriegs et al. (2006).

Statistical Analyses

To test the validity of clade hypotheses at various nodes of the phylogenetic tree, we used the method of Waddell et al. (2001). Assuming the sequence-based tree reconstruction as a prior hypothesis at any given node, a minimum of 3 integration sites against zero for alternative branching hypotheses are required for a significance level of P < 0.04.

Tissue Sampling

We isolated genomic DNA from tissue samples conserved in 95% ethanol from the following species using standard protocols (Sambrook et al. 1989)-Folivora: Bradypus tridactylus (pale-throated 3-toed sloth) and Choloepus hoffmanni (Hoffmann's 2-toed sloth); Vermilingua: Tamandua tetradactyla (southern tamandua), Myrmecophaga tridactyla (giant anteater), and Cyclopes didactylus (silky anteater); Dasypodinae: D. novemcinctus (9-banded armadillo) and Dasypus kappleri (greater long-nosed armadillo); Euphractinae: Euphractus sexcinctus (6-banded armadillo), Zaedyus pichiy (pichi), and Chaetophractus vellerosus (screaming hairy armadillo); Chlamyphorinae: Chlamyphorus truncatus (pink fairy armadillo); Tolypeutinae: Tolypeutes tricinctus (3-banded armadillo), Cabassous unicinctus (southern naked-tailed armadillo), and Priodontes maximus (giant armadillo); and Afrotheria (outgroup): Loxodonta africana (African savanna elephant) and Trichechus manatus (Caribbean manatee).

PCR Amplification

We initiated our screening with high-throughput PCR for all selected 283 loci, using DNA from *C. hoffmanni*, *M. tridactyla*, *D. novemcinctus*, *C. vellerosus*, *T. tricinctus*, and *T. manatus*. PCR primers for conserved, potentially informative loci were generated for amplificates not longer than 1 kb. Subsequently, we focused on 39 loci with detectable fragment size shifts (Schmitz et al. 2001) to amplify in the full species sampling. The following conditions were used for PCR amplifications: 4 min at 94 °C, 35 cycles of 30 s at 94 °C, 30 s at the primer-specific annealing temperature (supplementary table 1, Supplementary Material online), and 90 s at 72 °C. PCR reactions were finished with 5 min at 72 °C.

The pDrive Cloning Vector (Qiagen, Hilden, Germany) was utilized for ligation of PCR fragments purified from 1% agarose gels. TOP10 electrocompetent cells (Invitrogen, Groningen, The Netherlands) were used for electroporation. Sequencing was performed with standard M13 forward and reverse primers (supplementary table 1, Supplementary Material online) and the AmpliTaq FS Big Dye Terminator Kit (PE Biosystems, Foster City, CA).

Sequence Alignment and Phylogenetic Analyses

All amplified sequences were manually aligned for the final 39 selected loci. For detection and classification of additional random insertions, we used the RepeatMasker Server (Smit and Green, RepeatMasker at http://www. repeatmasker.org/cgi-bin/WEBRepeatMasker) and the CENSOR software tool (Kohany et al. 2006; http:// www.girinst.org/censor/index.php). Newly obtained sequences were submitted to the National Center for Biotechnology Information (NCBI) database (NCBI accession numbers EF465537–EF466007).

The raw concatenation of the 39 alignments yielded a combined data set of 28,869 sites for the 16 species. For phylogenetic reconstructions from these flanking regions of retroposed elements, we adopted a conservative approach. First, regions containing the retroposed elements were entirely removed from the concatenation to make subsequent phylogenetic inferences fully independent of the retroposed insertions. Then, ambiguously aligned sites and highly gapped/repeated regions were excluded, reducing the data set to 12,488 sites of which 5,530 were variable. Alignments are available upon request.

Phylogenetic analyses were conducted using maximum likelihood (ML) and Bayesian approaches. The best fitting model of sequence evolution was selected using Modeltest 3.7 (Posada and Crandall 1998). Among 56 competing models, the Akaike information criterion (Akaike 1974) identified the transversional model (TVM) with among-site rate heterogeneity modeled by a gamma distribution with 8 discrete categories ($+\Gamma_8$) as the best fitting model for our concatenated data set. The TVM model is a restriction of the most general, general time reversible (GTR) model. It allows different base compositions and different substitution rate probabilities for the 4 types of transversions, whereas the 2 types of transitions have the same rate.

The homogeneity of base composition across taxa was evaluated by the χ^2 test implemented in PAUP* 4.0b10 (Swofford 2002). This test did not reject the null hypothesis of overall base composition homogeneity (P = 0.64) in this A + T rich (62%) non-protein-coding data set. ML phylogenetic analyses were conducted with PAUP* through heuristic searches with tree bisection reconnection (TBR) branch swapping starting from a Neighbor-Joining (NJ) tree. ML model parameters were estimated with a successive approximation approach (Sullivan et al. 2005) using 3 rounds of optimization, starting from a NJ topology derived from ML distances obtained from a GTR + Γ_8 model $(\alpha = 0.5)$ and followed by ML heuristic searches by TBR branch swapping under the best fitting TVM + Γ_8 model. Reliability of nodes was estimated by bootstrap resampling with 500 pseudoreplicate data sets generated by the program SeqBoot of the PHYLIP package (Felsenstein 2001). The 500 corresponding ML heuristic searches were performed using PAUP* with the previously estimated ML parameters, NJ starting trees, and TBR branch swapping. They were parallelized on a Beowulf-class cluster supercomputer comprising 16 bi-processor Xeon 5160 Dual Core Intel computing nodes enabling simultaneous running of up to 64 processes. ML bootstrap percentages were obtained from the majority-rule consensus tree of the 500 bootstrap-generated trees using the latest version of Tree-Finder (Jobb et al. 2004).

Bayesian phylogenetic analyses of the concatenated flanking sequence data set were implemented using MrBayes 3.1.2 (Ronquist and Huelsenbeck 2003). Two independent runs of 4 incrementally heated Metropoliscoupled Markov chains Monte Carlo (MCMCMC) were launched from random starting points. We used Dirichlet priors for base frequencies (1,1,1,1) and for GTR parameters (1,1,1,1,1) scaled to the G–T substitution rate, a Uniform (0.05, 50.00) prior for the Γ distribution, and an Exponential (10.0) prior for branch lengths. The 4 MCMCMC were simultaneously run for 1,000,000 generations using the program default priors as starting values for the GTR + Γ_8 model parameters. Parameters and trees were sampled every 100 generations, and the convergence of MCMCMC was checked by monitoring values of the marginal likelihood and model parameters through generations. Bayesian posterior probabilities (PP) were obtained from the 50% majority-rule consensus of the 2 × 9,000 trees sampled at stationarity after removing the first 2 × 1,000 trees as the "burn-in" stage.

Results

From \sim 7,400 investigated loci, 283 were selected for high-throughput PCRs in representatives of all major xenarthran branches and an afrotherian outgroup (Materials and Methods). A total of 244 loci did not contain xenarthranspecific RGCs or were not amplifiable in all representative lineages. For the remaining 39 informative loci, we extended PCR screening and sequencing of retroposed elements and their flanking regions including all available species (Materials and Methods). These 39 informative loci contained 35 phylogenetically informative retroposed element markers, 14 autapomorphic retroposon insertions, and 39 phylogenetically informative indels other than retroposons (fig. 1; supplementary table 2, Supplementary Material online). For most of the diagnostic retroposed elements, including spliceosomal U snRNAs (retropseudogenes generated by a variant L1 retroposition mechanism [Schmitz et al. 2004]), we also detected the short DR and the singular target sites at orthologous positions in species devoid of the retroposons. Moreover, examination of the phylogenetically conserved random indels occurring in flanking regions was restricted to those with a minimum length of 3 nt.

The ML tree reconstructed from the concatenated 12,488 sites of retroposon flanking sequences resulted in a well-resolved phylogeny (fig. 2). All but 4 nodes were supported by bootstrap percentages (BPs) of 100 and PPs of 1.0. This ML topology strongly supported xenarthran monophyly and their division into the 3 major groups Folivora (sloths), Vermilingua (anteaters), and Cingulata (armadillos). Anteaters and sloths were firmly grouped together into Pilosa (BP = 99; PP = 1.0). Within anteaters, *Myrmecophaga* and *Tamandua* were sister groups (BP = 100; PP = 1.0) to the exclusion of the early diverging *Cyclopes*.

Within armadillos, phylogenetic analyses unambiguously supported (BP = 100; PP = 1.0) the respective monophyly of the 3 previously identified subfamilies: Dasypodinae, Euphractinae, and Tolypeutinae. Among Euphractinae, the sister group of *Zaedyus* and *Chaetophractus* was strongly supported (BP = 99; PP = 1.0), whereas relationships were less clear within Tolypeutinae, where *Tolypeutes* and *Cabassous* are sister groups (BP = 81; PP = 0.99). Finally, *Chlamyphorus*, representing the previously unsampled subfamily Chlamyphorinae, showed clear phylogenetic affinities with the clade Tolypeutinae (BP = 93; PP = 1.0).



FIG. 1.—Presence/absence of retroposed elements in the 39 investigated loci. Small letters denote more than one diagnostic RGC per locus. (*a*) Retroposon presence/absence matrix. DAS-SINEs are highlighted in gray. Plus indicates the presence of a retroposed element, minus indicates its absence, and question mark denotes that PCR amplification did not work. (*b*) Indel markers in the 39 investigated loci. Insertions are marked with black, deletions with white boxes. One indicates the presence of an indel, zero indicates the original locus without the indel. Question mark denotes that the PCR amplification failed. For outgroup comparisons, we selected the afrotherians *Loxodonta africana* or *Trichechus manatus*.

In the 39 informative presence/absence loci examined, we found independent support for 10 internal branches of the ML tree (fig. 3)—1) Xenarthra: 5 independent insertions of retroposed elements (U6, U5-L1MB2-L1MB4-chimera, L1MB4, L1MB5, L1ME3A) confirm the monophyly of xenarthrans. The corresponding elements were clearly absent in Afrotherian outgroup species. 2) Pilosa: 1 LINE element (L1MC2) and 2 indels were present in all 5 representative species of Bradypodidae, Megalonychidae, and Myrmecophagidae and were absent in all species of Dasypodidae, thus providing evidence for the monophyly of Pilosans. 3) Folivora: a diagnostic 27 nt deletion was detected in the Bradypodidae and Megalonychidae species that was not present in other Pilosa. 4) Myrmecophagidae (Vermilingua): we found 1 LINE element (L1MB8) and a 24 nt deletion in all 3 representatives of Myrmecophagidae. 5) Myrmecophaginae: 4 indels present in Myrmecophaga and *Tamandua* indicate a common ancestry of these species to the exclusion of *Cyclopes*. 6) Dasypodidae (Cingulata): 17 diagnostic retroposons (10× DAS-Ia, 3× DAS-IIa, 2× DAS-IIb, U2, L1-like element) and 3 indels were found in the Dasypodidae that were absent in all pilosans. 7) Dasypodinae: 8 retroposons (3 × DAS-Ia, 2 × DAS-IIa, DAS-III), U2, L1MB2) and 16 indels were present in only the 2 investigated *Dasypus* species. 8) Euphractinae + Chlamyphorinae + Tolypeutinae: we detected 1 DAS-IIb element and 3 indels in the Euphractinae, Chlamyphorinae, and Tolypeutinae species that were not present in Dasypodinae. 9) Euphractinae: 1 DAS-IIb insertion and 7 indels were present only in the Euphractinae species. 10) *C. vellerosus* + *Z. pichiy*: a DAS-III3b element insertion and a 19 nt diagnostic deletion support the grouping of *Chaetophractus* and *Zaedyus*.

Discussion

SINE Activity in Xenarthran Genomes

Of the identified retroposed elements, 36 were Cingulataspecific DAS-SINEs. As our search strategies were based predominantly on available sequence information from *D. novemcinctus*, most diagnostic insertions were found in the lineage leading to this species. Diagnostic insertions on other xenarthran branches were found as additional





FIG. 2.—Sequence- and retroposon-based phylogenetic trees. (*a*) ML phylogram based on 12,488 concatenated sites from flanking sequence of the 39 RGCs containing loci under the TVM + Γ_8 model. ML estimates of model parameters were A = 0.32, C = 0.19, G = 0.19, T = 0.30; A \leftrightarrow C = 0.96, A \leftrightarrow G = 3.46, A \leftrightarrow T = 0.77, C \leftrightarrow G = 1.18, C \leftrightarrow T = 3.46, G \leftrightarrow T = 1.00; \alpha = 1.73. Values at nodes represent bootstrap percentages/Bayesian PP, respectively. (*b*) Phylogenetic tree based on discrete retroposon insertions. Filled circles mark the insertion of DAS-SINEs; open circles mark the insertion of other retroposed elements such as LINEs or snRNAs. All identifying numbers refer to those used in figures 1 and 3 and supplementary tables 1 and 2 (Supplementary Material online). The monophyly of Xenarthra, Cingulata, and Dasypodinae are significantly supported.



FIG. 3.—Independent phylogenetic evidence from retroposon insertion and indel events mapped on the ML tree of Xenarthra inferred from flanking sequences. Filled circles mark the insertion of DAS-SINEs; open circles mark the insertion of other retroposed elements such as LINEs or snRNAs. Triangles mark the presence of diagnostic indels. All insertions of retroposed elements and indels are assigned identifying number and letter codes that are used in figure 1 and 2*b* and supplementary tables 1 and 2 (Supplementary Material online). Indel #25' (right brackets) contradicts our ML tree by merging *Chlamyphorus truncatus* and *Cabassous unicinctus*. However, this indel occurs in a variable region (see text for details).

insertions in the investigated loci. Unfortunately sufficient sequence data to enable similar searches are not yet available for other xenarthran species. This is expressed in the limited significant phylogenetic support from discrete retroposon data (fig. 2b). The generation of large-scale genomic sequence information (e.g., by Bacterial Artificial Chromosome or whole-genome shotgun libraries) will provide a multidirectional source for additional lineagespecific informative loci in the future. The DAS-SINEs preferentially investigated in this study spread in characteristic waves of activity that correspond to the evolutionary scenario proposed by Churakov et al. (2005). Simple monomeric DAS-SINEs like DAS-Ia were active in ancient armadillo lineages, and complex elements like DAS-III3b were active in species of some terminal branches.

The armadillo-specific distribution of DAS-SINEs was confirmed by screening all available genomic sequences of *C. hoffmanni* using a DAS-SINE–specific library (RepeatMasker). Not a single copy of DAS-like sequences was detected. Interestingly, we found no other lineage-specific SINE elements for other xenarthrans than those for Cingulata (but see also Nishihara et al. 2007). A more elaborated search (e.g., for SINE-specific promoter boxes) will be necessary to fill this gap. However, we did find lineage-specific L1-like and L1MB2 elements in Cingulata and L1MC2 and L1MB8 elements in Pilosa.

Improved Resolution of Xenarthran Phylogeny

Reconstructing phylogenetic relationships among living xenarthran genera has been the subject of previous molecular studies using a combination of 3 nuclear and 2 mitochondrial genes (Delsuc et al. 2002, 2003). These studies, including all extant genera except Chlamyphorus, resulted in a comprehensive phylogenetic framework with only 2 remaining uncertainties within armadillos. The present study represents the first ever evidence from retroposed elements for resolving xenarthran phylogeny. Also, the concatenation of flanking sequences offered the first large-scale, non-protein-coding data set from which the evolutionary history of all living xenarthran genera could be inferred. The results from the present work independently corroborated previous studies and provided strong additional support for most internal branches within Xenarthra including the areas of uncertainty.

Indeed, previous molecular data failed to resolve relationships within the subfamilies Tolypeutinae and Euphractinae. Contradictory results were obtained from mitochondrial versus nuclear genes resulting in low statistical support values when analyzing their concatenation (Delsuc et al. 2003). These difficult to resolve trifurcations were interpreted as evidence for the occurrence of 2 radiation events possibly triggered by environmental changes (Delsuc et al. 2004). Consensus network analyses of the phylogenetically conflicting signals present in the combination of 2 mitochondrial and 3 nuclear genes (Delsuc and Douzery, 2008) revealed that 2 of the 3 competing hypotheses coexisted with almost equal support within each subfamily: *Chaetophractus* + *Euphractus* versus *Chaetophractus* + *Zaedyus* within Euphractinae and *Cabassous* + *Priodontes* versus *Cabassous* + *Tolypeutes* within Tolypeutinae.

Phylogenetic analyses of the more than 12-kb data set of non-protein-coding DNA provided unambiguous support for the grouping of the genera *Chaetophractus* and *Zaedyus* within Euphractinae (fig. 2*a*). Furthermore, *C. vellerosus* and *Z. pichiy* shared a diagnostic DAS-III3b element insertion and a 19-nt-long deletion providing independent evidence for their direct common ancestry to the exclusion of *E. sexcinctus* (fig. 3; supplementary figure 1, Supplementary Material online). This result is in contradiction to a cladistic study of craniodental characters favoring a sister group relationship between *Chaetophractus* and *Euphractus* supported by 5 synapomorphies (Gaudin and Wible 2006), but it agrees with the relationships proposed by Abrantes and Bergqvist (2006) based on additional postcranial evidence.

Our results appeared less clear-cut for the relationships within Tolypeutinae as no diagnostic retroposed element insertions or indel events were found (fig. 3). Sequencebased phylogenetic analyses, nevertheless, favor the monophyly of *Cabassous* and *Tolypeutes* to the exclusion of *Priodontes* (fig. 2*a*). This is also in contrast to the morphological evidence for grouping *Cabassous* with *Priodontes* (Abrantes and Bergqvist 2006; Gaudin and Wible 2006). However, in this case, it is possible that morphologically based cladistic reconstructions might have been confounded by the highly derived morphology of *Tolypeutes* relative to other members of Tolypeutinae. Obviously, more work is needed to understand the reasons of these discrepancies between molecular and morphological characters.

The present work provides the first molecular data to resolve the phylogenetic position of the armadillo subfamily Chlamyphorinae (fairy armadillos) within Cingulata. In contrast with cladistic analyses of morphological characters (Abrantes and Bergqvist 2006; Gaudin and Wible 2006), the probabilistic phylogenetic analyses of the flanking region data set strongly supported a close relationship of C. truncatus with Tolypeutinae instead of Euphractinae (fig. 2a). However, no diagnostic RGC was found to confirm this unexpected result (fig. 3). Future extensive wetbench approaches focusing on retroposons in these and other lineages will help to strengthen the sequence-based relationships and balance biased in silico methods that depend on available sequence sources. However, for the present work only very little tissue was available from the rare species. This constrained our analyses to preferentially in silico methods instead of the tissue intensive wet-bench approaches. The inclusion of the second, morphologically divergent, species of fairy armadillo (C. retusus) in future sequence-based phylogenetic analyses might also help to confirm the position of Chlamyphorinae within the xenarthran tree by breaking up the branch leading to C. truncatus.

Of all the retroposed elements and indels identified, only one indel (#25') contradicted our ML tree. This 19 nt deletion was found in C. truncatus and C. unicinctus and not in other species. It is nevertheless worth noting that in this region, 2 species of *Dasypus* also contained a 57 nt deletion encompassing the one shared by Chlamyphorus and Cabassous. The occurrence of 2 embedded deletions in the same region casts doubt on the potential phylogenetic significance of these deletion events that might be homoplasic. We also cannot fully exclude neither lineage sorting nor paralogy effects, but the probability of comparing paralogs is very low with respect to the highly conserved regions compared (this conservation encompasses xenarthrans and afrotherians up to human) and the definitively singular fragments in PCRs of all compared species. The contradiction presented by this random indel emphasizes the necessity to search for additional support by more reliable retroposon insertions using experimental lineage-specific search strategies as outlined in Endoh and Okada (1986) or by random sequencing and subsequent Zoo-PCR.

The results of the present study make Xenarthra the first major placental group to have its phylogenetic history fully resolved at the genus level by using multiple independent molecular and genomic evidences. These results provide a robust phylogenetic framework for future evolutionary studies of comparative genomics, taxonomy, cytology, and development in this mammalian order. It might also be useful for defining conservation priorities for its most endangered species (da Fonseca and Aguiar 2004).

Supplementary Material

Supplementary figure 1 and tables 1 and 2 are available at *Molecular Biology and Evolution* online (http://www.mbe.oxfordjournals.org/).

Acknowledgments

We thank Pablo Beldomenico, Benoit de Thoisy, Martin S. Fischer, Jorgue Omar García, Eric Hansen, Rodney L. Honeycutt, Jan Ole Kriegs, Jean-François Mauffrey, Bernhard Neurohr, Rodolpho Rearte, Sylvia Rolle, Erica Taube, Jean-Christophe Vié, and Heike Weber for providing us with tissue samples. We are grateful to François Catzeflis for providing access to the Mammalian Tissue Collection of the Institut des Sciences de l'Evolution de Montpellier. We also thank Marsha Bundman for editorial assistance. Two anonymous reviewers and the associated editor provided helpful suggestions on the manuscript. This work was supported by the Deutsche Forschungsgemeinschaft (SCHM 1469 to J.S. and J.B.). This is contribution ISEM 2007-080 of the Institut des Sciences de l'Evolution de Montpellier (UMR5554-CNRS).

Literature Cited

Abrantes EAL, Bergqvist LP. 2006. Proposta filogenética para os Dasyposidae (Mammalia: cingulata). In: Gallo V, Brito PM, Silva HMA, Figueiredo FJ, editors. Paleontologia de Vertebrados: Grandes Temas e Contribuições Científicas. Rio De Janeiro: Interciência Ltda. p. 261–274.

- Akaike H. 1974. A new look at the statistical model identification. IEEE Trans. Autom. Contr. AC-19. 716–723.
- Amrine-Madsen H, Koepfli KP, Wayne RK, Springer MS. 2003. A new phylogenetic marker, apolipoprotein B, provides compelling evidence for eutherian relationships. Mol Phylogenet Evol. 28:225–240.
- Anderson RP, Handley CO. 2001. A new species of three-toed sloth (Mammalia: xenarthra) from Panamá, with a review of the genus *Bradypus*. Proc Biol Soc Wash. 114:1–33.
- Borodulina OR, Kramerov DA. 2005. PCR-based approach to SINE isolation: simple and complex SINEs. Gene. 349: 197–205.
- Cetica PD, Solari AJ, Merani MS, De Rosas JC, Burgos MH. 1998. Evolutionary sperm morphology and morphometry in armadillos. J Submicrosc Cytol Pathol. 30:309–314.
- Churakov G, Smit AF, Brosius J, Schmitz J. 2005. A novel abundant family of retroposed elements (DAS-SINEs) in the nine-banded armadillo (*Dasypus novemcinctus*). Mol Biol Evol. 22:886–893.
- Cuéllar E. 2001. The tatujeikurajoyava (*Chlamyphorus retusus*) in the Izozog communities of the Bolivian Gran Chaco. Edentata. 4:14–16.
- da Fonseca GAB, Aguiar JM. 2004. The 2004 edentate species assessment workshop. Edentata. 6:1–26.
- Delsuc F, Catzeflis FM, Stanhope MJ, Douzery EJP. 2001. The evolution of armadillos, anteaters and sloths depicted by nuclear and mitochondrial phylogenies: implications for the status of the enigmatic fossil Eurotamandua. Proc R Soc B. 268:1605–1615.
- Delsuc F, Douzery EJP. Forthcoming. Recent advances and future prospects in xenarthran molecular phylogenetics. In: Loughry WJ, Vizcaíno SF, editors. Biology of the Xenarthra. Gainesville (FL): University Press of Florida.
- Delsuc F, Scally M, Madsen O, Stanhope MJ, De Jong WW, Catzeflis FM, Springer MS, Douzery EJP. 2002. Molecular phylogeny of living xenarthrans and the impact of character and taxon sampling on the placental tree rooting. Mol Biol Evol. 19:1656–1671.
- Delsuc F, Stanhope MJ, Douzery EJP. 2003. Molecular systematics of armadillos (Xenarthra, Dasypodidae): contribution of maximum likelihood and Bayesian analyses of mitochondrial and nuclear genes. Mol Phylogenet Evol. 28: 261–275.
- Delsuc F, Vizcaíno SF, Douzery EJP. 2004. Influence of tertiary paleoenvironmental changes on the diversification of South American mammals: a relaxed molecular clock study within xenarthrans. BMC Evol Biol. 4:11.
- Endoh H, Okada N. 1986. Total DNA transcription in vitro: a procedure to detect highly repetitive and transcribable sequences with tRNA-like structures. Proc Natl Acad Sci USA. 83:251–255.
- Engelmann GF. 1985. The phylogeny of the Xenarthra. In: Montgomery GG, editor. The evolution and ecology of armadillos, sloths and vermilinguas. Washington (DC): Smithsonian Institution Press. p. 51–63.
- Felsenstein J. 2001. PHYLIP (PHYLogeny Inference Package). Version 3.66. Department of Genome Sciences. Seattle (WA): University of Washington.
- Gaudin TJ, Branham DG. 1998. The phylogeny of Myrmecophagidae (Mammalia, Xenarthra, Vermilingua) and the relationship of Eurotamandua to the Vermilingua. J Mamm Evol. 5:237–265.
- Gaudin TJ, Wible JR. 2006. The phylogeny of living and extinct armadillos (Mammalia, Xenarthra, Cingulata): a craniodental

analysis. In: Carrano MT, Gaudin TJ, Blob RW, Wible JR, editors. Amniote paleobiology: perspectives on the evolution of mammals, birds and reptiles. Chicago (IL): University of Chicago Press. p. 153–198.

- Greenwood AD, Castresana J, Feldmaier-Fuchs G, Pääbo S. 2001. A molecular phylogeny of two extinct sloths. Mol Phylogenet Evol. 18:94–103.
- Höss M, Dilling A, Currant A, Pääbo S. 1996. Molecular phylogeny of the extinct ground sloth Mylodon darwinii. Proc Natl Acad Sci USA. 93:181–185.
- Jobb G, von Haeseler A, Strimmer K. 2004. TREEFINDER: a powerful graphical analysis environment for molecular phylogenetics. BMC Evol Biol. 4:18.
- Kohany O, Gentles AJ, Hankus L, Jurka J. 2006. Annotation, submission and screening of repetitive elements in Repbase: repbaseSubmitter and Censor. BMC Bioinformatics. 7:474.
- Kriegs JO, Churakov G, Kiefmann M, Jordan U, Brosius J, Schmitz J. 2006. Retroposed elements as archives for the evolutionary history of placental mammals. PLoS Biol. 4:e91.
- Madsen O, Scally M, Douady CJ, Kao DJ, DeBry RW, Adkins R, Amrine HM, Stanhope MJ, de Jong WW, Springer MS. 2001. Parallel adaptive radiations in two major clades of placental mammals. Nature. 409:610–614.
- McKenna MC. 1975. Toward a phylogenetic classification of the Mammalia. In: Luckett WP, Szalay FS, editors. Phylogeny of the Primates. New York: Plenum Press. p. 21–46.
- McKenna MC, Bell SK. 1997. Classification of Mammals above the species level. New York: Columbia University Press.
- Murphy WJ, Eizirik E, Johnson WE, Zhang YP, Ryder OA, O'Brien SJ. 2001. Molecular phylogenetics and the origins of placental mammals. Nature. 409:614–618.
- Murphy WJ, Eizirik E, O'Brien SJ, et al. (11 co-authors). 2001. Resolution of the early placental mammal radiation using Bayesian phylogenetics. Science. 294:2348–2351.
- Murphy WJ, Pringle TH, Crider TA, Springer MS, Miller W. 2007. Using genomic data to unravel the root of the placental mammal phylogeny. Genome Res. 17:413–421.
- Nikaido M, Rooney AP, Okada N. 1999. Phylogenetic relationships among cetartiodactyls based on insertions of short and long interpersed elements: hippopotamuses are the closest extant relatives of whales. Proc Natl Acad Sci USA. 96:10261–10266.
- Nikolaev S, Montoya-Burgos JI, Margulies EH, Program NC, Rougemont J, Nyffeler B, Antonarakis SE. 2007. Early history of mammals is elucidated with the ENCODE multiple species sequencing data. PLoS Genet. 3:e2.
- Nishihara H, Hasegawa M, Okada N. 2006. Pegasoferae, an unexpected mammalian clade revealed by tracking ancient retroposon insertions. Proc Natl Acad Sci USA. 103: 9929–9934.
- Nishihara H, Kuno S, Nikaido M, Okada N. 2007. MyrSINEs: a novel SINE family in the anteater genomes. Gene. 400:98–103.
- Okada N. 1991. SINEs: short interspersed repeated elements of the eukaryotic genome. Trends Ecol Evol. 6:358–361.
- Patterson B, Pascual R. 1972. The fossil mammal fauna of South America. In: Keast A, Erk FC, Glass BP, editors. Evolution, mammals and southern continents. Albany (NY): State University of New York Press. p. 247–309.
- Patterson B, Segall W, Turnbull WD. 1989. The ear region in xenarthrans (= Edentata, Mammalia). Part I. Cingulates. Fieldiana. Geology. 18:1–46.
- Patterson B, Segall W, Turnbull WD, Gaudin TJ. 1992. The ear region in xenarthrans (= Edentata, Mammalia). Part II. Sloths, anteaters, palaeanodonts, and a miscellany. Fieldiana. Geology. 24:1–79.
- Posada D, Crandall KA. 1998. MODELTEST: testing the model of DNA substitution. Bioinformatics. 14:817–818.

- Ray DA, Xing J, Salem AH, Batzer MA. 2006. SINEs of a nearly perfect character. Syst Biol. 55:928–935.
- Reiss KZ. 1997. Myology of the feeding apparatus of myrmecophagid anteaters (Xenarthra: Myrmecophagidae). J Mamm Evol. 4:87–117.
- Rokas A, Holland PW. 2000. Rare genomic changes as a tool for phylogenetics. Trends Ecol Evol. 15:454–459.
- Ronquist F, Huelsenbeck JP. 2003. MrBayes 3: bayesian phylogenetic inference under mixed models. Bioinformatics. 19: 1572–1574.
- Roos C, Schmitz J, Zischler H. 2004. Primate jumping genes elucidate strepsirrhine phylogeny. Proc Natl Acad Sci USA. 101:10650–10654.
- Salem AH, Ray DA, Xing J, Callinan PA, Myers JS, Hedges DJ, Garber RK, Witherspoon DJ, Jorde LB, Batzer MA. 2003. Alu elements and hominid phylogenetics. Proc Natl Acad Sci USA. 100:12787–12791.
- Sambrook J, Fritsch EF, Maniatis T. 1989. Molecular cloning: a laboratory manual, 2nd ed. Cold Spring Harbor (NY): Cold Spring Harbor Laboratory Press.
- Schmitz J, Churakov G, Zischler H, Brosius J. 2004. A novel class of mammalian-specific tailless retropseudogenes. Genome Res. 14:1911–1915.
- Schmitz J, Ohme M, Zischler H. 2001. SINE insertions in cladistic analyses and the phylogenetic affiliations of *Tarsius bancanus* to other primates. Genetics. 157:777–784.
- Shedlock AM, Okada N. 2000. SINE insertions: powerful tools for molecular systematics. Bioessays. 22:148–160.
- Steel MA, Penny D. 2005. Maximum parsimony and the phylogenetic information in multistate characters. In: Albert V, editor. Parsimony, phylogeny and genomics. Oxford: Oxford University Press. p. 163–178.

- Sullivan J, Abdo Z, Joyce P, Swofford DL. 2005. Evaluating the performance of a successive-approximations approach to parameter optimization in maximum-likelihood phylogeny estimation. Mol Biol Evol. 22:1386–1392.
- Superina M. 2006. New information on population declines in pink fairy armadillos. Edentata. 7:48–50.
- Swofford DL. 2002. PAUP*. Phylogenetic analysis using parsimony (and other methods). Version 4.0b10. Sunderland (MA): Sinauer Associates.
- Van Dijk MAM, Paradis E, Catzeflis F, De Jong WW. 1999. The virtues of gaps: xenarthran (Edentate) monophyly supported by a unique deletion in alpha A-crystallin. Syst Biol. 48: 94–106.
- Vizcaíno SF. 1995. Identificación específica de las "mulitas", género *Dasypus* L. (Mammalia, Dasypodidae), del noroeste argentino. Descripción de una nueva espécie. Mastozoología Neotropical. 2:5–13.
- Waddell PJ, Kishino J, Ota R. 2001. A phylogenetic foundation for comparative mammalian genomics. Genome Inform. 12: 141–154.
- Webb SD. 1985. The interrelationships of tree sloths and ground sloths. In: Montgomery GG, editor. The evolution and ecology of armadillos, sloths and vermilinguas. Washington (DC): Smithsonian Institution Press. p. 105–112.
- Wetzel RM. 1985. The identification and distribution of recent Xenarthra (= Edentata). In: Montgomery GG, editor. The evolution and ecology of armadillos, sloths and vermilinguas. Washington (DC): Smithsonian Institution Press. p. 5–21.

Norihiro Okada, Associate Editor

Accepted September 16, 2007